



# The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts

## Citation

Inouye, Sharon K., Cyrus M. Kosar, Douglas Tommet, Eva M. Schmitt, Margaret R. Puelle, Jane S. Saczynski, Edward R. Marcantonio, and Richard N. Jones. 2014. "The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts." *Ann Intern Med* 160 (8) (April 15): 526. doi:10.7326/m13-1927.

## Published Version

doi:10.7326/M13-1927

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33749831>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Published in final edited form as:

*Ann Intern Med.* 2014 April 15; 160(8): 526–533. doi:10.7326/M13-1927.

## The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts

Sharon K. Inouye, M.D., MPH<sup>1,2</sup>, Cyrus M. Kosar, MA<sup>2</sup>, Douglas Tommet, MS, MPH<sup>2</sup>, Eva M. Schmitt, PhD<sup>2</sup>, Margaret R. Puelle<sup>2</sup>, Jane S. Saczynski, PhD<sup>3</sup>, Edward R. Marcantonio, M.D., SM<sup>#1</sup>, and Richard N. Jones, ScD<sup>#2,4</sup>

<sup>1</sup>Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

<sup>2</sup>Aging Brain Center, Institute for Aging Research, Hebrew SeniorLife

<sup>3</sup>Division of Geriatric Medicine and Meyers Primary Care Institute, University of Massachusetts Medical School, Worcester, MA

<sup>4</sup>Departments of Psychiatry and Human Behavior and Neurology, Brown University Warren Alpert Medical School, Providence RI

# These authors contributed equally to this work.

### Abstract

**Background**—Quantifying the severity of delirium is essential to advance clinical care through improved understanding of delirium impact, prognosis, pathophysiology, and response to treatment.

**Objectives**—To develop and validate a new delirium severity measure (CAM-S) based on the Confusion Assessment Method

**Design**—Validation analysis in two independent cohorts.

**Setting**—Three academic medical centers

**Patients**—First cohort included 300 patients age ≥ 70 years scheduled for major surgery; second included 919 medicine admissions age ≥ 70 years.

**Measurements**—A 4-item short form and 10-item long-form were developed. The association of the maximal CAM-S score during hospitalization with hospital and post-hospital outcomes related to delirium was evaluated.

**Results**—CAM-S scores demonstrated strong associations with all clinical outcomes examined, with significant gradients across severity categories in adjusted analyses, adding substantively to delirium diagnoses alone. Representative results included adjusted mean length of stay (LOS), which increased across levels of CAM-S short form severity from 6.5 (95% confidence interval, CI, 6.2-6.9) to 12.7 days (95% CI, 11.2-14.3) ( $P_{\text{trend}} < 0.001$ ). Comparable results across

increasing levels of the CAM-S long form severity were 5.6 (95% CI, 5.1-6.1) to 11.9 days (95% CI, 10.8-12.9) ( $P_{\text{trend}} < 0.001$ ). Representative results for the composite outcome of adjusted relative risk of death or nursing home residence at 90 days increased across levels of CAM-S short form severity from 1.0 (referent) to 2.5 (95% CI, 1.9-3.3) ( $P_{\text{trend}} < 0.001$ ). Comparable results for the CAM-S long form severity were 1.0 (referent) to 2.5 (95% CI, 1.6-3.7) ( $P_{\text{trend}} < 0.001$ ).

**Limitations**—Data on clinical outcomes were drawn from an older dataset involving patients age 70 years and older.

**Conclusions**—CAM-S provides a new delirium severity measure with strong psychometric properties and strong associations with important clinical outcomes.

### Keywords

Delirium; Confusion Assessment Method; elderly; validation; predictive validity

## Introduction

Delirium, a common, serious, costly, and potentially preventable condition for older persons, has been identified as the leading complication of hospitalization for older persons (1, 2). Given the associated hospital mortality rates of 25-33% (3,4) and annual healthcare costs of >\$182 billion (2011 USD) (5), delirium has assumed increasing attention as a public health and patient safety priority (2,6,7). The Confusion Assessment Method (CAM) is a standardized, validated measure (8) that has gained widespread use in screening for delirium (2,9,10). To date, the CAM has been used in over 4000 original articles, and translated into 14 languages. While previous studies have used an additive score of CAM features as a severity indicator (11, 12), the validity of this approach has not been examined. Extending the CAM as a severity measure would greatly enhance its clinical value.

A CAM-based severity measure would have substantial utility in both clinical and research applications. Clinically, a measure to track delirium severity would be useful to follow response to delirium treatment and management interventions across clinical settings. The measure could also be useful to demonstrate the graded impact of delirium on healthcare delivery, such as clinical care staffing and costs. Measurement of delirium severity is essential to understand its clinical course and recovery; thus, such a measure could serve as a primary outcome for prognostic studies and treatment trials. The measure would facilitate studies of pathophysiology, where quantifying the level of delirium and its change over time may clarify mechanistic relationships. Importantly, the availability of a standardized measure would facilitate comparison across studies.

Thus, we developed a new CAM-based scoring system for delirium severity, called the CAM-S. The specific aims were: (1) to create the CAM-S scoring system (including a short form based on the 4-item CAM algorithm and a long form based on the 10-item CAM instrument) (8); (2) to evaluate its distribution, test properties, and inter-rater reliability; (3) to examine how it performs in persons with dementia; and (4) to determine its association with clinical outcomes likely to be related to delirium severity. We hypothesized that a valid severity measure should be a strong independent predictor of adverse outcomes associated

with delirium, including prolonged hospital stay, functional and cognitive decline, nursing home placement, death, and high healthcare costs.

## Methods

### Study Samples

The two samples were prospective cohort studies with consecutive sampling, described previously. The Successful Aging after Elective Surgery (SAGES) study, currently ongoing, provided the first sample (13). Potential participants were consecutive patients scheduled for elective major noncardiac surgery at two Harvard affiliated hospitals from June 10, 2010 to March 29, 2012. A total of 951 patients met initial eligibility criteria: age  $\geq 70$  years, scheduled for major orthopedic, vascular, or general surgical procedures under general or regional anesthesia, anticipated to have at least a 2-day hospital stay, and living within 50 miles. Of these, 446 patients were ineligible because of delirium ( $n=2$ ), clinically documented dementia ( $n=14$ ), hospitalization within 3 months ( $n=78$ ), terminal condition ( $n=12$ ), legal blindness or severe deafness ( $n=3$ ), non-English speaking ( $n=80$ ), or other reasons, primarily related to being unreachable by phone or unable to communicate verbally ( $n=257$ ). An additional 205 patients refused participation, and eligibility could not be determined. The final sample included 300 patients, which is a subgroup of the total planned sample.

Project Recovery, described previously (5, 11, 14), provided the second sample. Potential participants were consecutive patients admitted to the medicine service at Yale New Haven Hospital from March 25, 1995 to March 18, 1998. A total of 2,434 patients met eligibility criteria: age  $\geq 70$  years, no delirium on admission but intermediate to high risk for delirium. Of these, 1,265 patients were excluded because of inability to participate in interviews ( $n=298$ ); [e.g., nonverbal due to profound dementia ( $n=154$ ), language barrier ( $n=92$ ), severe aphasia ( $n=38$ ), respiratory isolation ( $n=12$ ), and intubation ( $n=2$ )]; coma or terminal illness ( $n=69$ ); hospital stay of  $< 48$  hours ( $n=219$ ); prior enrollment ( $n=324$ ); or other reasons ( $n=355$ ). An additional 250 patients refused enrollment. The final sample included 919 participants.

Assessment of clinical outcomes was conducted by research assistants (interviewers and chart abstractors) who were blinded to the CAM delirium status of the patients; the CAM ratings were conducted by a separate hospital-based team of research assistants. The screening and recruitment of the participants in each cohort was conducted before the CAM was performed. For the SAGES study, informed consent for participation was obtained from all patients according to procedures approved by the institutional review boards (IRBs) at the Harvard-affiliated hospitals. For Project Recovery, informed consent was obtained from the patients or, for those with substantial cognitive impairment, from a proxy as approved by the Yale IRB.

### Delirium Assessment

In both cohorts, delirium was rated using the CAM, which was scored daily during hospitalization by experienced research assistants based on observations made during a

standardized interview, including a sleep questionnaire and brief cognitive tests, which took about 10-15 minutes. Study interviewers underwent intensive training and standardization (15).

### Development and Scoring of the CAM-S

The CAM-S is intended to be used *in addition to* the original CAM algorithm; that is, the CAM-S will not yield a delirium diagnosis, only a means to quantify the intensity of delirium symptoms observed at the bedside. These symptoms can be present in persons both with and without delirium. We created a short-form and long-form of the CAM-S scoring system. The short form was based on the four features from the CAM diagnostic algorithm (7) which can be rated at the bedside: acute onset or symptom fluctuation, inattention, disorganized thinking, and altered level of consciousness. Each symptom of delirium--except fluctuation--was rated as absent (0), mild (1) or marked (2). Acute onset or fluctuation was rated as absent (0) or present (1). The sum of these ratings yielded a CAM-S short form severity score ranging from 0 to 7 (7 = most severe). The long form was based on the 10 features from the full CAM instrument (8): acute onset or symptom fluctuation, inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation, and sleep-wake cycle disturbance. Each symptom was rated 0-2, except for acute onset or fluctuation, as previously described. The sum of these ratings yielded a CAM-S long form score from 0 to 19 (19 = most severe). Features scored as "uncertain" did not contribute to the severity score. Uncertain ratings were present for one or more items in only 13/1456 (< 1%) CAM-S short form items and 38/1456 (< 3%) CAM-S long form items.

### Inter-Rater Reliability

To assess inter-rater reliability, a total of 73 paired CAM-S ratings (14 delirious and 59 non-delirious patients) have been conducted on a quarterly basis in the SAGES study to date, with 2 observers rating each patient simultaneously in a blinded fashion.

### Evaluation of Convergent Agreement

To demonstrate convergent agreement, closely related measures should be highly correlated. Convergent agreement was assessed by examining the correlation of daily CAM-S scores with concurrent measures of confusion and cognitive functioning completed daily during hospitalization. In SAGES, convergent agreement was assessed by comparing daily CAM-S scores with a brief cognitive screening test administered to the patients (SAGES COG screen) (13) and a global rating of confusion (scored 0-10, higher worse) rated by the interviewers daily. In Project Recovery, convergent agreement was assessed by comparing daily CAM-S scores with the concurrent Mini-Mental State Examination administered to the patients (scored 0-30, higher better, licensed from PAR, Inc.) (16) and a visual analog scale for confusion (scored 0-100, higher worse) rated by the interviewers.

### Association with Clinical Outcomes

The association with clinical outcomes was assessed in Project Recovery, where data collection was complete; these clinical outcomes are not yet available in SAGES. Hospital

outcomes included length of stay and nursing home placement abstracted from the medical record, and hospital costs from the hospital's billing database. The costs are for hospital services that were submitted on UB-92 hospital billing forms to Medicare Part A, which typically represent about 50-60% of hospital charges and do not include professional fees and services (Medicare Part B). Functional decline defined as decline by 2 or more points on the 14-point Activities of Daily Living (ADL) score (equivalent to a decline in 1 ADL on a standard 7-point scale) between baseline and discharge (17, 18, 19). Cognitive decline defined as decline by 2 or more points on the 30-point Mini-Mental State Examination (MMSE) score between baseline and discharge (18, 20). Post-hospital outcomes included death within 90 days which counted hospital deaths, determined from medical records, Medicare Part A and Social Security databases, National Death Index, and death certificates (5, 21). Nursing home residence at 90 days and post-discharge healthcare costs per day for 90 days were obtained from Medicare Part A data. To avoid bias, the hierarchical outcome of either death or nursing home residence was used, since patients who die can no longer be placed in a nursing home. Functional decline at 30 days was defined by decline by 2 or more points on the 14-point ADL score (equivalent to a decline in 1 ADL on a standard 7-point scale) between baseline and 30-day follow-up interview; no 90-day interview was conducted.

### Statistical Analyses

[See Appendix for further details on validation analyses]. For all analyses, one measure per patient (the most severe CAM-S score during hospitalization) was used; the only exception was convergent agreement where all observations were used for purposes of daily comparison. Inter-rater reliability was estimated with overall agreement on exact scores and intraclass correlation coefficients. Convergent agreement was estimated by Pearson's correlation with daily measures of cognitive functioning in the hospital. To rule out bias due to selecting more than one observation per person, we repeated our analyses selecting only one pair of daily CAM-S and cognitive functioning scores from each person at random. We also verified fulfillment of the linearity assumption (assumed by the correlation coefficient) by comparing total variance explained by linear and polynomial models.

To enhance interpretability for analyses related to clinical outcomes, we divided the CAM-S scores into four categories defined empirically based on score distributions. Poisson regression was used to calculate adjusted relative risks (RR) for outcomes related to nursing home, death, cognitive and functional decline. For analyses relevant to new nursing home placement, patients not alive or already living in a nursing home before the relevant time frame were excluded. Finally, we modeled hospital length of stay (LOS) and Medicare costs using log-gamma regression (24). Predictive margins (25) were obtained and presented as adjusted mean length of stay and costs. Because delirium severity was strongly related to death, we calculated costs per day to avoid bias in the post-hospital analyses. Post-hospital costs per day were not available for 88 patients with missing cost data, which included deaths and enrollees in health maintenance organizations. Linear trend tests were used to evaluate the exposure-response relationship between delirium severity and clinical outcomes. All models were adjusted for age, sex, race, APACHE II score, Charlson comorbidity index, and dementia. In addition, all models except for functional decline were

also adjusted for baseline impairment in activities of daily living. For all models, we verified the robustness of our parameter estimates to potentially influential (e.g., high leverage) observations by repeating models excluding these observations. Graphic displays of the raw data on maximal CAM-S scores by clinical outcomes are presented in Appendix Figures 1-9.

To examine the additional contribution of delirium severity beyond delirium diagnosis, all outcomes were reexamined in analyses stratified by delirium status. CAM-S scores were modeled as continuous measures, then presented in tertiles based on the within-group CAM-S distributions to enhance interpretability. The tertiles were intended to have approximately equal-sized groups, but were limited by the constraints of integer-based scores. Linear trend tests were repeated within each stratum and for the overall sample.

All analyses were conducted with Stata Version 13 (26). The funding sources, including grants from the National Institute on Aging, had no role in the study design, conduct, or reporting.

## Results

Characteristics of the patients are shown in Table 1. Both samples included older adults (mean ages 77 and 80 years, respectively) with a predominance of women (55% and 60%, respectively). However, compared with SAGES, Project Recovery had higher baseline comorbidity (Charlson 2.9 vs. 1.0), higher levels of functional and cognitive impairment at baseline, and higher rates of possible dementia at baseline (13% vs. 2%). While clinically documented dementia was an exclusionary criterion for SAGES, early dementia was present in 2% according to a Clinical Consensus panel convened after enrollment. In the SAGES sample, delirium developed during hospitalization in 68/300 (23%), with missing CAM ratings in 16/1456 observations (1%) on both short and long forms. In Project Recovery, delirium developed during hospitalization in 115/919 (13%), with missing CAM ratings in < 19/5202 observations (< 0.5%) on both short and long forms. There were no adverse events associated with administering the CAM.

For SAGES, the mean (SD) CAM-S scores were 0.68 (SD 1.22) for the short-form and 2.00 (SD 2.50) for the long-form; for Project Recovery, the mean (SD) CAM-S scores were 0.33 (SD 0.94) for the short-form and 1.90 (SD 2.20) for the long-form. The distribution of CAM-S in delirious and non-delirious patients is shown in Figure 1. For the CAM-S short form, a 3-point difference in the median separates the delirious and non-delirious groups in both samples, with no overlap in their interquartile ranges (25-75<sup>th</sup> percentile values). The same relationship is demonstrated for the CAM-S long-form, with a 6-point difference in median values and no overlap in interquartile ranges. In stratified analyses, the CAM-S severity scores (medians and interquartile ranges) were appropriately separated between delirious and non-delirious groups in the overall sample and in those with dementia, supporting construct validity. Since SAGES excluded dementia patients, these analyses were conducted only in Project Recovery. The short-form shows a 4-point difference in median values for severity with no overlap in interquartile ranges between delirious and non-



delirious patients with dementia. Similarly, the long-form shows a 5-point difference in median values for severity with no overlap in interquartile ranges.

In 73 paired observations, including 19% delirious patients, the mean (SD) CAM-S scores were 1.24 (SD 1.65) for the short-form and 3.00 (SD 3.55) for the long-form. The overall agreement for CAM-S short form scoring was 98% with an intraclass correlation coefficient of 0.92, indicating high agreement. For the CAM-S long form, the overall agreement was 97%, with an intraclass correlation coefficient of 0.88, also demonstrating high agreement.

The CAM-S demonstrated moderate to high convergent agreement with other daily cognitive measures. In SAGES, the CAM-S agreement with was high with the daily confusion rating ( $r=0.78$ ,  $0.80$  with short and long forms, respectively) and brief cognitive screen ( $r=0.62$ ,  $0.72$ ). In Project Recovery, the CAM-S agreement was moderate to high with the daily confusion rating ( $r=0.45$ ,  $0.64$ ) and MMSE ( $r=0.41$ ,  $0.64$ ). Similar results were obtained utilizing repeated random samples of one observation per patient, and using polynomial rather than linear models.

Table 2 demonstrates the association of CAM-S with hospital outcomes, where the associations demonstrated significant trends for all outcomes. Length of hospital stay increases across each CAM-S short form severity category from an adjusted mean of 6.5 days for no delirium symptoms to 12.7 days with high severity. A similar gradient is seen for the CAM-S long form, increasing across categories from an adjusted mean of 5.6 to 11.9 days. Hospital costs (all in 1995 US dollars) increase across CAM-S categories from an adjusted mean of \$5100 for no delirium symptoms to \$13,200 for severe delirium. A similar gradient is seen for the CAM-S long form, increasing from an adjusted mean of \$4200 for no delirium symptoms to \$11,400 for severe delirium. The adjusted relative risks of new nursing home placement were 1.0, 1.4, 2.1 and 2.5 ( $p\text{-trend} < 0.001$ ) across CAM-S short form categories, and 1.0, 1.4, 2.3 and 3.9 ( $p\text{-trend} < 0.001$ ) across CAM-S long form categories. The proportion of patients with functional decline between baseline and discharge increased across both CAM-S short form (from 36% to 68%) and long form (from 25% to 61%) categories. Cognitive decline between baseline and discharge also increased across severity categories (from 16% to 65% for short form and 10% to 50% for long form).

The association of CAM-S with post-hospital outcomes is shown in Table 3, where the associations demonstrated significant trends for all outcomes. The cumulative rates of death within 90 days increases across each CAM-S short form category from 7% for no delirium symptoms to 27% with high severity ( $aRR = 3.3$ ). The gradient for the CAM-S long form ranges from a cumulative rate of death of 7% to 22% ( $aRR = 2.3$ ) from the lowest to highest categories. Finally, rates of death or nursing home residence at 90 days increase across each CAM-S short form category from 15% to 51% ( $aRR = 2.5$ ) and from 13% to 48% ( $aRR = 2.5$ ) for the long form. In those available for the one-month follow-up interview, functional decline at 30 days increased across severity categories from 29% to 52% ( $aRR 1.6$ ) for the short form; and from 20% to 52% ( $aRR 2.3$ ) for the long form.

Appendix Tables 1-2 repeat the analyses from Tables 2 and 3, stratified by delirium diagnosis. These analyses reveal the additional contribution of CAM-S within strata defined



by the presence or absence of delirium. Within each subgroup in Appendix Tables 1-2, poorer clinical outcomes were seen with higher levels of CAM-S scores. These trends are all statistically significant in the non-delirious subgroup. With the small sample sizes in the delirium subgroups, however, the trends did not achieve statistical significance for many outcomes in this subgroup.

In Appendix Figures 1-9, the contribution of CAM-S beyond delirium diagnosis is suggested by the spread of points within each subgroup, indicating higher CAM-S scores with worse adverse outcomes regardless of delirium status.

## Discussion

This study provides evidence for the usefulness and validity of a new delirium severity measure, the CAM-S. Since this measure is based on the CAM, which is already widely used, it poses distinct advantages for future work. This study demonstrates that the CAM-S has good psychometric properties, high inter-rater reliability, and importantly, strong association with clinical outcomes related to delirium. Both the short and long forms of the CAM-S demonstrate strong psychometric properties (See also Appendix) and each holds unique advantages. The short form (5 minute completion time) is quicker and simpler to rate based on the CAM diagnostic algorithm alone; however, the long form (10 minute completion time) provides a broader range of severity scores in both delirium and non-delirium groups.

Strengths of this study include the rigorous prospective validation of the CAM-S in two independent samples. While the many differences between the two study populations might be viewed as a limitation, in fact, their disparate nature lends strong support that the CAM-S will work well in different populations and under different conditions, supporting generalizability of the findings. The use of state-of-the-art methods for delirium assessment, high quality data with relatively little missing data, and the broad range of clinical outcomes for comparison enhanced the validation process. The rich clinical database with long-term follow-up enabled detailed assessment of the impact of CAM-S on important clinical outcomes.

Several important caveats about this study are worthy of mention. First, the age of the Project Recovery data may be viewed as a limitation. The relatively long lengths of hospital stay and low hospital costs are reflective of hospital care in 1995-1998; however, we were primarily interested in the comparison of these outcomes between our severity groups. Thus, internal validity of the comparisons is paramount and the absolute values were less important. Another potential limitation is that the patients in both cohorts are age 70 and older, and the performance of the CAM-S will require future validation in younger adults. In addition, another limitation is that there may be inherent dependencies between CAM-S and adverse outcomes; for instance, patients with longer lengths of stay may have had higher CAM-S scores because of more opportunities for measurement. Finally, for accurate rating of the CAM-S, brief but formal cognitive testing of the patients should be conducted, which will require training and standardization of the staff along with some additional staff time.

Since the CAM is well-known and widely used, the CAM-S may pose distinct logistical advantages over existing delirium severity measures and fills an important gap in the applicability of the CAM. Unlike the Delirium Rating Scale (27, 28), a clinician rater is not required for the CAM-S, rather well-trained research assistants can reliably conduct the assessments. Compared with the Memorial Delirium Assessment Scale (29), the CAM-S short form is quicker and simpler to use. In addition, the CAM-S short form has the important advantage of being relatively unbiased with respect to rating the severity of either hyperactive and hypoactive forms of delirium, since the features rated are present in both forms. All delirium severity measures to date have been limited by the over-representation of hyperactive features compared with hypoactive features (that is, hyperactive features like agitation and hallucinations contribute more to the total severity score than hypoactive features like psychomotor retardation). While not directly examined in the present study, the severity of hypoactive delirium, which is the more common type among older persons, may not be captured adequately with these instruments. This imbalance holds important implications for clinical trials targeting management of delirium. If a treatment converts patients from a hyperactive to hypoactive delirium, the treatment may be rated as efficacious if the outcome measure does not adequately capture the severity of the hypoactive delirium. The fact that the results of many clinical trials for delirium have shown discrepant results (2, 7, 30), with many not demonstrating benefits for (or even worsening of) clinical outcomes, may be a direct consequence of these measurement limitations. Delirium severity is a complex and multifaceted construct, and weighing the relative contributions of different symptom categories (e.g., cognitive versus behavioral, hyperactive versus hypoactive) can pose unique challenges. It is possible that separate severity ratings for these different symptom subgroups may be needed. Finally, examining the relative and combined contributions of delirium severity, duration, and recurrence to outcomes is essential to better define the clinical impact of delirium. While beyond the scope of the present study, these are important areas for future investigation.

The CAM-S provides a new scoring method with strong psychometric properties to add to the armamentarium of delirium measures. This measure may serve as a primary outcome for clinical trials and studies of the pathophysiology or prognosis of delirium. It is hoped that the availability of this measure will serve to facilitate critically-needed studies of delirium and its outcomes, and ultimately lead to improved quality of life for older persons and their families.

## Acknowledgments

The authors gratefully acknowledge the patients, families, physicians, and research staff who participated in the SAGES and Project Recovery studies and who made this study possible.

This work is dedicated to the memory of Joshua Bryan Inouye Helfand and Bradley Yoshio Inouye.

The Confusion Assessment Method algorithm and instrument is copyrighted to the Hospital Elder Life Program, LLC <[www.hospitalelderlifeprogram.org](http://www.hospitalelderlifeprogram.org)>

The analytic code and documentation for this study is available upon request.

**Grant Funding:** Supported by Grants No. P01AG031720 (SKI), K07AG041835 (SKI), R01AG030618 (ERM), and K24AG035075 (ERM) from the National Institute on Aging. Dr. Inouye holds the Milton and Shirley F. Levy

Family Chair. Dr. Saczynski's time was supported in part by Grant No. K01AG033643 from the National Institutes on Aging. The funding sources had no role in the design, conduct, or reporting of this study.

## Appendix

### Validation Analyses: Factor Analysis

Unidimensionality and model fit were assessed with permuted parallel analysis (22) and confirmatory factor analysis in the SAGES sample. Internal reliability was assessed by Cronbach's alpha and McDonald's omega (23).

Parallel and confirmatory factor analyses demonstrate that the CAM-S is a unidimensional measure. The single factor model demonstrates good fit (confirmatory fit index = 0.99 and root mean squared error of approximation = 0.04). Good internal reliability of the CAM-S is demonstrated by the Cronbach's alpha = 0.73, 0.73 and McDonald's omega = 0.92, 0.90 for the CAM-S short and long-forms, respectively.

**Appendix Table 1**  
**Association of CAM-S with Hospital Outcomes Stratified by Delirium Diagnosis\***

Hospital Outcome	Length of Stay	Hospital Costs	New Nursing Home Placement		Functional Decline		Cognitive Decline	
	Adj. Mean, Days (95% CI)	Adj. Mean (95% CI), \$ × 1000	n/N (%)	Adj. Mean Probability (95% CI)	n/N (%)	Adj. Mean Probability (95% CI)	n/N (%)	Adj. Mean Probability (95% CI)
CAM-S Short Form Severity Rating								
Non-Delirious Group	N=804	N=804	n=757	N=794	N=793			
Scores of 0 (n=598)	6.5 (6.2, 6.8)	5.1 (4.8, 5.4)	61/571 (11)	10.6 (8.3, 13.0)	212/589 (36)	36.3 (32.6, 40.1)	93/589 (16)	16.2 (13.3, 19.1)
Scores of 1 (n=91)	8.4 (8.0, 8.9)	6.8 (6.3, 7.3)	14/85 (16)	17.2 (13.7, 20.6)	46/90 (51)	48.2 (43.6, 52.7)	23/89 (26)	21.1 (17.3, 24.8)
Scores of 2-3 (n=115)	11.3 (10.1, 12.4)	9.6 (8.3, 10.8)	26/101 (26)	25.4 (18.1, 32.6)	69/115 (60)	60.6 (52.5, 68.6)	27/115 (23)	25.1 (17.8, 32.4)
<i>P</i> <sub>trend</sub>	<.001	<.001	---	<.001	---	<.001	---	0.02
Delirious Group	N=115	N=115	N=94	N=114	N=109			
Scores of 2 (n=15) <sup>†</sup>	11.8 (9.0, 14.6)	10.3 (7.0, 13.6)	3/12 (25)	26.3 (11.2, 41.5)	11/15 (73)	62.6 (46.3, 79.0)	7/13 (54)	50.9 (34.2, 67.6)
Scores of 3-4 (n=64)	12.4 (10.6, 14.2)	12.0 (9.7, 14.4)	21/57 (37)	35.4 (25.9, 44.8)	38/63 (60)	65.9 (56.5, 75.3)	36/61 (59)	59.9 (50.0, 69.8)
Scores of 5-7 (n=36)	15.1 (11.8, 18.4)	16.4 (11.5, 21.3)	11/25 (44)	46.7 (30.1, 63.4)	31/36 (86)	80.8 (69.5, 92.0)	28/35 (80)	79.6 (68.0, 91.2)
<i>P</i> <sub>trend</sub>	0.26	0.03	---	0.22	---	0.08	---	0.02
CAM-S Long Form Severity Rating								
Non-Delirious Group	N=804	N=804	n=757	N=794	N=793			
Scores of 0-1 (n=205)	5.8 (5.4, 6.2)	4.5 (4.2, 4.9)	13/198 (7)	6.3 (4.2, 8.4)	50/202 (25)	30.2 (25.6, 34.8)	20/202 (10)	13.5 (10.2, 16.9)
Scores of 2 (n=288)	6.9 (6.6, 7.2)	5.6 (5.2, 5.9)	25/278 (9)	9.4 (7.3, 11.6)	115/284 (40)	38.2 (34.6, 41.8)	53/282 (19)	16.4 (13.6, 19.2)
Scores of 3-10 (n=311)	8.9 (8.3, 9.5)	7.2 (6.6, 7.9)	63/281 (22)	22.2 (18.2, 26.2)	162/308 (53)	51.1 (46.3, 56.0)	70/309 (23)	22.4 (18.3, 26.6)
<i>P</i> <sub>trend</sub>	<.001	<.001	---	<.001	---	<.001	---	<.01
Delirious Group	N=115	N=115	N=94	N=114	N=109			
Scores of 4-6 (n=41)	11.7 (9.5, 13.9)	10.8 (8.0, 13.6)	9/35 (26)	27.0 (15.0, 39.0)	27/40 (68)	67.0 (54.8, 79.2)	18/38 (47)	50.8 (37.4, 64.1)
Scores of 7-10 (n=42)	13.0 (11.2, 14.8)	12.6 (10.2, 15.0)	16/38 (42)	42.7 (33.3, 52.2)	27/42 (64)	66.7 (57.6, 75.9)	29/40 (73)	66.6 (57.7, 75.6)

Hospital Outcome	Length of Stay		Hospital Costs		New Nursing Home Placement		Functional Decline		Cognitive Decline	
	Adj. Mean, Days (95% CI)		Adj. Mean (95% CI), \$ × 1000		n/N (%)	Adj. Mean Probability (95% CI)		n/N (%)	Adj. Mean Probability (95% CI)	
Scores of 11-19 (n=32)	15.3 (11.7, 18.8)		17.2 (11.6, 22.8)		10/21 (48)	44.4 (27.3, 61.6)		26/32 (81)	78.7 (67.0, 90.3)	
$P_{\text{trend}}$	0.29	<.05			---	0.13	---	0.19	---	<.01

Functional decline defined as decline by 2 or more points on the 14-point Activities of Daily Living (ADL) score (equivalent to 1 ADL on the standard 7-point scale) between baseline and discharge. Cognitive decline defined as decline by 2 or more points on the 30-point Mini-Mental State Examination (MMSE) score between baseline and discharge. Adj=adjusted; CI=confidence interval. All models are adjusted for age, sex, nonwhite race, Apache II score, Charlson comorbidity index, and baseline dementia. All models except the model for functional decline were also adjusted for baseline impairment in activities of daily living. Nursing home analyses excluded 54 patients residing in nursing home at baseline and 14 who died during hospitalization. The functional decline analyses excluded 11 patients with missing ADL information at discharge; the cognitive decline analyses excluded 17 patients with missing MMSE values at discharge.

\* Analyses conducted in Project Recovery. The maximal CAM-S score during each patient's hospitalization was used in all analyses. Adjusted mean probabilities are predictive margins obtained from logistic regression.

<sup>†</sup>This delirious group (with CAM-S scores of 2) was based on clinical adjudication and chart review. Sensitivity analyses excluding this group reveal similar results.

Appendix Table 2  
Association of CAM-S with Post-Hospital Outcomes Stratified by Delirium Diagnosis\*

Post-Hospital Outcome	Death at 90 Days		Cost Per Day, First 90 Days		Nursing Home Residence or Death at 90 Days		Functional Decline at 30 Days	
	n/N (%)	Adj. Mean Probability (95% CI)	Adj. Mean (95% CI), \$	n/N (%)	Adj. Mean Probability (95% CI)	n/N (%)	Adj. Mean Probability (95% CI)	
CAM-S Short Form Severity Rating								
Non-Delirious Group	N=804		N=732		N=735		N=650	
Scores of 0 (n=598)	39/598 (7)	6.9 (5.0, 8.8)	109.8 (95.6, 124.1)	81/544 (15)	15.3 (12.4, 18.2)	143/497 (29)	29.1 (25.3, 32.9)	
Scores of 1 (n=91)	14/91 (15)	11.4 (8.9, 14.0)	151.0 (126.6, 175.4)	27/82 (33)	27.4 (23.2, 31.6)	29/70 (41)	38.1 (33.2, 43.0)	
Scores of 2-3 (n=115)	18/115 (16)	17.0 (11.0, 23.0)	204.2 (147.3, 261.0)	41/109 (38)	39.6 (31.5, 47.7)	36/83 (43)	44.1 (34.6, 53.6)	
<i>P</i> <sub>trend</sub>	---	<.001	<.001	---	<.001	---	<.01	
Delirious Group	N=115		N=99		N=109		N=62	
Scores of 2 (n=15) <sup>†</sup>	2/15 (13)	14.5 (3.6, 25.3)	174.2 (112.4, 235.9)	8/14 (57)	41.1 (27.0, 55.1)	5/10 (50)	46.6 (25.2, 67.9)	
Scores of 3-4 (n=64)	15/64 (23)	21.9 (14.1, 29.8)	182.3 (148.0, 216.6)	27/61 (44)	49.8 (41.0, 58.6)	17/35 (49)	49.8 (36.7, 62.8)	
Scores of 5-7 (n=36)	13/36 (36)	38.3 (24.3, 52.3)	226.8 (160.6, 292.9)	21/34 (62)	58.5 (44.6, 72.3)	11/17 (65)	64.3 (44.9, 83.7)	
<i>P</i> <sub>trend</sub>	---	0.02	0.21	---	0.32	---	0.32	
CAM-S Long Form Severity Rating								
Non-Delirious Group	N=804		N=732		N=735		N=650	
Scores of 0-1 (n=205)	14/205 (7)	5.5 (3.5, 7.4)	91.5 (76.1, 107.0)	24/181 (13)	11.2 (8.2, 14.2)	35/175 (20)	22.8 (18.2, 27.3)	
Scores of 2 (n=288)	20/288 (7)	8.3 (6.3, 10.4)	115.1 (100.6, 129.7)	37/263 (14)	16.2 (13.3, 19.1)	67/238 (28)	27.7 (24.0, 31.4)	
Scores of 3-10 (n=311)	37/311 (12)	11.5 (8.5, 14.5)	164.7 (133.6, 195.7)	88/291 (30)	29.6 (25.1, 34.1)	106/237 (45)	43.1 (37.7, 48.6)	
<i>P</i> <sub>trend</sub>	---	<.01	<.001	---	<.001	---	<.001	
Delirious Group	N=115		N=99		N=109		N=62	
Scores of 4-6 (n=41)	8/41 (20)	16.2 (7.4, 24.9)	161.7 (118.8, 204.6)	16/38 (42)	37.9 (26.6, 49.1)	12/26 (46)	46.9 (30.1, 63.6)	
Scores of 7-10 (n=42)	8/42 (19)	23.8 (16.1, 31.5)	208.7 (166.3, 251.1)	20/40 (50)	56.3 (48.2, 64.5)	12/21 (57)	53.1 (40.4, 65.8)	

Post-Hospital Outcome	Death at 90 Days		Cost Per Day, First 90 Days		Nursing Home Residence or Death at 90 Days		Functional Decline at 30 Days	
	n/N (%)	Adj. Mean Probability (95% CI)	Adj. Mean (95% CI), \$	n/N (%)	Adj. Mean Probability (95% CI)	n/N (%)	Adj. Mean Probability (95% CI)	n/N (%)
Scores of 11-19 (n=32)	14/32 (44)	41.8 (27.5, 56.1)	216.1 (151.3, 281.0)	20/31 (65)	61.5 (47.7, 75.4)	9/15 (60)	64.4 (44.8, 84.1)	
<i>P</i> <sub>trend</sub>	---	<.01	0.18	---	0.06	---	0.22	

In those who were available for telephone follow up interviews at one-month (N=650), functional decline defined as decline by 2 or more points on the 14-point Activities of Daily Living (ADL) score (equivalent to 1 ADL on the standard 7-point scale) between baseline and 30 days. Adj=adjusted; CI=confidence interval. Deaths within 90 days include all in-hospital deaths. All models are adjusted for age, sex, nonwhite race, Apache II score, Charlson comorbidity index, and baseline dementia. All models except the model for functional decline were also adjusted for baseline impairment in activities of daily living. Medicare data (nursing home and costs) were missing for 75 patients (including those receiving care in health maintenance organizations). The cost per day analyses additionally excluded 13 patients who died during hospitalization.

\* Analyses conducted in Project Recovery. The maximal CAM-S score during each patient's hospitalization was used in all analyses. Adjusted mean probabilities are predictive margins obtained from logistic regression. Adjusted mean costs and length of stay are predictive margins obtained from log-gamma regression.

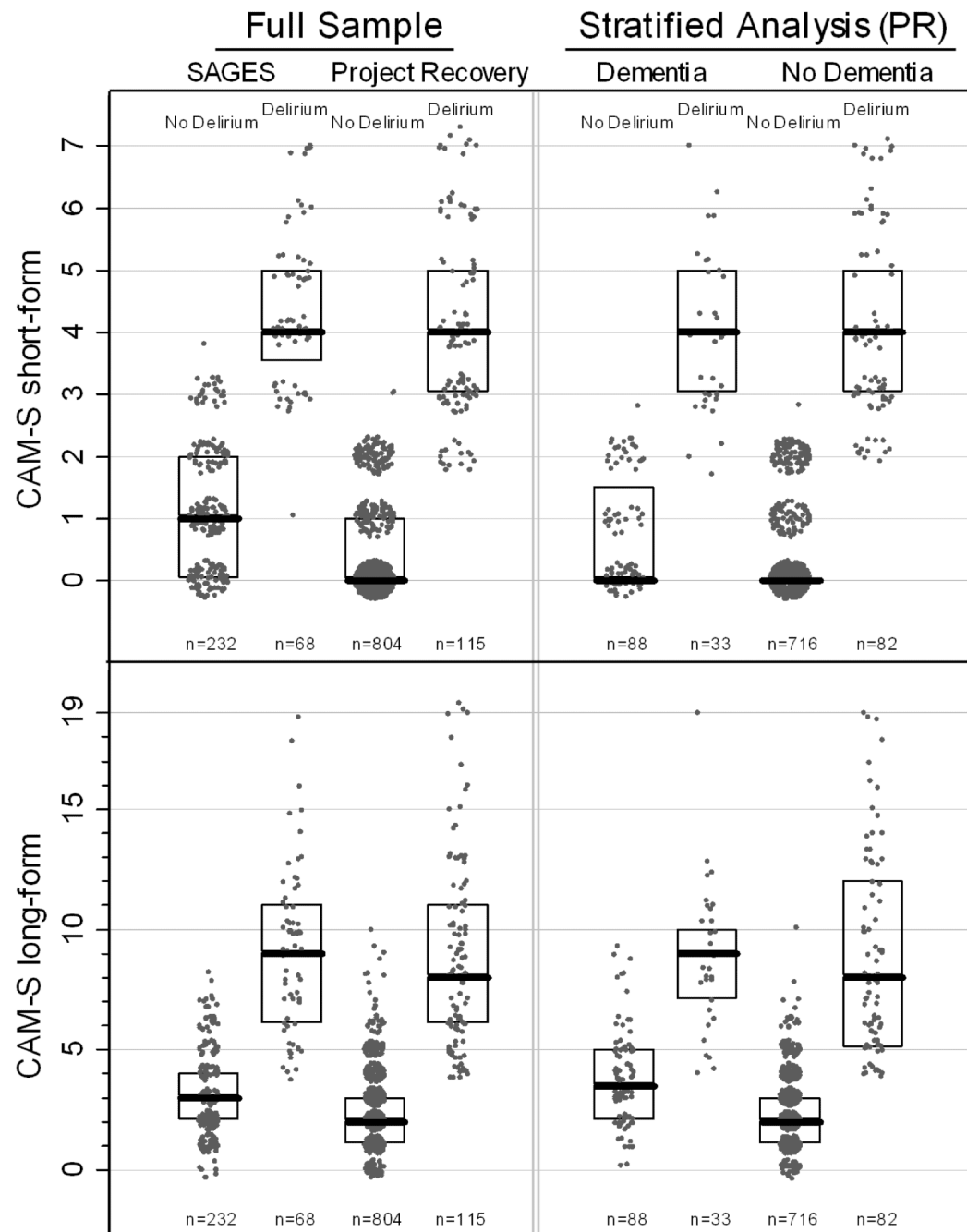
† This delirious group (with CAM-S scores of 2) was based on clinical adjudication and chart review. Sensitivity analyses excluding this group reveal similar results.



## References

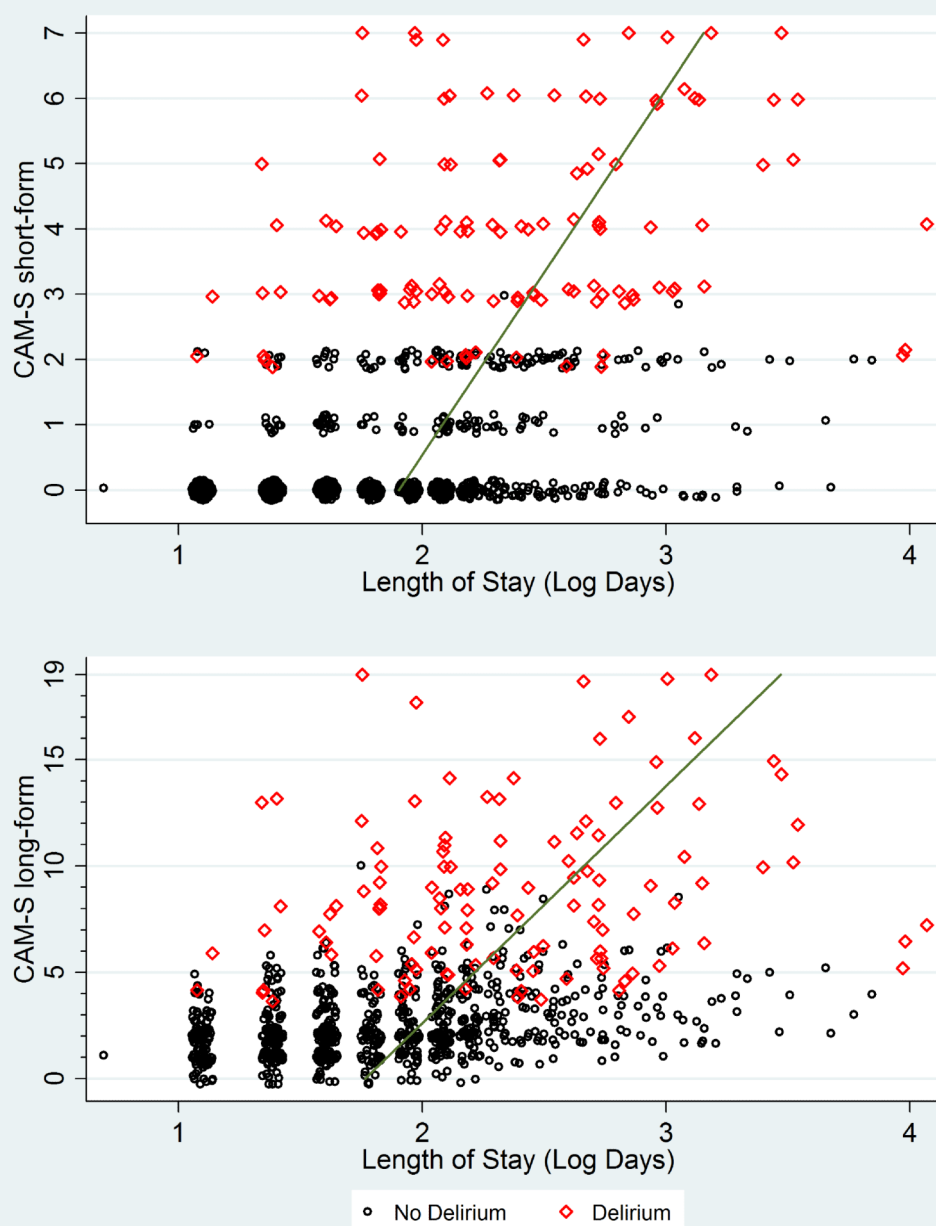
1. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2013 Epub ahead of print. [PMID: 23992774].
2. National Institute for Health and Clinical Excellence. Delirium: diagnosis, prevention and management. CG103. National Institute for Health and Clinical Excellence; London: 2010.
3. Inouye SK. Delirium in older persons. *N Engl J Med*. 2006; 354:1157–65. PMID: 16540616. [PubMed: 16540616]
4. Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA*. 2010; 304:443–51. PMID: 20664045. [PubMed: 20664045]
5. Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK. One-year health care costs associated with delirium in the elderly population. *Arch Intern Med*. 2008; 168:27–32. PMID: 18195192. [PubMed: 18195192]
6. Wachter, RM. Understanding patient safety. 2nd Edition. McGraw-Hill Medical; New York, NY: 2012.
7. Greer N, Rossom R, Anderson P, MacDonald R, Tacklind J, Rutks I, et al. Delirium: Screening, Prevention, and Diagnosis--A Systematic Review of the Evidence. VA Evidence-based Synthesis Program Reports. 2011 PMID: 22206108.
8. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med*. 1990; 113:941–8. PMID: 2240918. [PubMed: 2240918]
9. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. *J Am Geriatr Soc*. 2008; 56:823–30. PMID: 18384586. [PubMed: 18384586]
10. Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. *JAMA*. 2010; 304:779–86. PMID: 20716741. [PubMed: 20716741]
11. Inouye SK, Bogardus ST Jr, Charpentier PA, Leo-Summers L, Acampora D, Holford TR, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med*. 1999; 340:669–76. PMID: 10053175. [PubMed: 10053175]
12. Milisen K, Foreman MD, Abraham IL, De Geest S, Godderis J, Vandermeulen E, et al. A nurse-led interdisciplinary intervention program for delirium in elderly hip-fracture patients. *J Am Geriatr Soc*. 2001; 49:523–32. PMID: 11380743. [PubMed: 11380743]
13. Schmitt EM, Marcantonio ER, Alsop DC, Jones RN, Rogers SO Jr, Fong TG, et al. Novel risk markers and long-term outcomes of delirium: the successful aging after elective surgery (SAGES) study design and methods. *J Am Med Dir Assoc*. 2012; 13:818 e1–10. PMID: 22999782. [PubMed: 22999782]
14. Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST Jr, Leslie DL, Agostini JV. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. *J Am Geriatr Soc*. 2005; 53:312–8. PMID: 15673358. [PubMed: 15673358]
15. Inouye, SK. The Confusion Assessment Method (CAM): Training Manual and Coding Guide. Yale University School of Medicine; New Haven: 2003. Available at: [www.hospitalelderlifeprogram.org](http://www.hospitalelderlifeprogram.org)
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975; 12:189–98. PMID: 1202204. [PubMed: 1202204]
17. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA: the journal of the American Medical Association*. 1963; 185:914–9. PMID: 14044222. [PubMed: 14044222]
18. Inouye SK, Bogardus ST Jr, Baker DI, Leo-Summers L, Cooney LM Jr. The Hospital Elder Life Program: a model of care to prevent cognitive and functional decline in older hospitalized patients. Hospital Elder Life Program. *J Am Geriatr Soc*. 2000; 48:1697–706. PMID: 11129764. [PubMed: 11129764]

19. Covinsky KE, Palmer RM, Fortinsky RH, Counsell SR, Stewart AL, Kresevic D, et al. Loss of independence in activities of daily living in older adults hospitalized with medical illnesses: increased vulnerability with age. *J Am Geriatr Soc.* 2003; 51:451–8. PMID: 12657063. [PubMed: 12657063]
20. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc.* 1992; 40:922–35. PMID: 1512391. [PubMed: 1512391]
21. Leslie DL, Zhang Y, Bogardus ST, Holford TR, Leo-Summers LS, Inouye SK. Consequences of preventing delirium in hospitalized older adults on nursing home costs. *J Am Geriatr Soc.* 2005; 53:405–9. PMID: 15743281. [PubMed: 15743281]
22. Buja A, Eyuboglu N. Remarks on parallel analysis. *Multivariate behavioral research.* 1992; 27:509–40.
23. McDonald, RP. *Test Theory: A Unified Treatment.* Erlbaum, editor. Psychology Press; Mahway, NJ: 1999.
24. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *Journal of Health Economics.* 2001; 20:461–494. [PubMed: 11469231]
25. Graubard BI, Korn EL. Predictive margins with survey data. *Biometrics.* 1999; 55:652–9. [PubMed: 11318229]
26. StataCorp. *Stata Statistical Software: Release 13.* StataCorp LP; College Station, TX: 2013.
27. Trzepacz PT, Baker RW, Greenhouse J. A symptom rating scale for delirium. *Psychiatry Res.* 1988; 23:89–97. PMID: 3363018. [PubMed: 3363018]
28. Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. *J Neuropsychiatry Clin Neurosci.* 2001; 13:229–42. PMID: 11449030. [PubMed: 11449030]
29. Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S. The Memorial Delirium Assessment Scale. *J Pain Symptom Manage.* 1997; 13:128–37. PMID: 9114631. [PubMed: 9114631]
30. Lacasse H, Perreault MM, Williamson DR. Systematic review of antipsychotics for the treatment of hospital-associated delirium in medically or surgically ill patients. *Ann Pharmacother.* 2006; 40:1966–73. PMID: 17047137. [PubMed: 17047137]



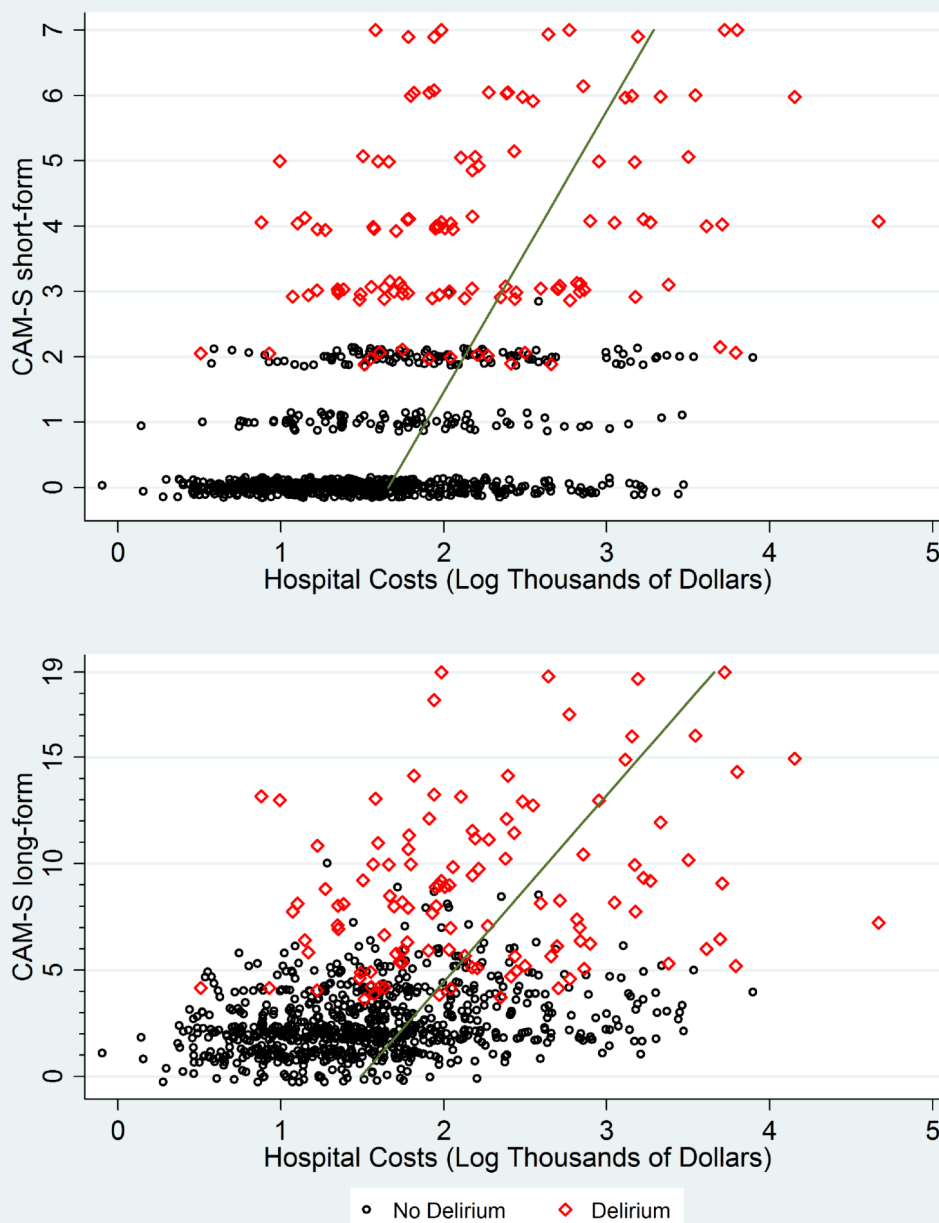
**Figure 1. Distribution of CAM-S by Delirium Status in Total Sample and Stratified by Dementia Status**

\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and 75<sup>th</sup> percentile all had the same value. The stratified analyses by dementia group were conducted in Project Recovery (PR).



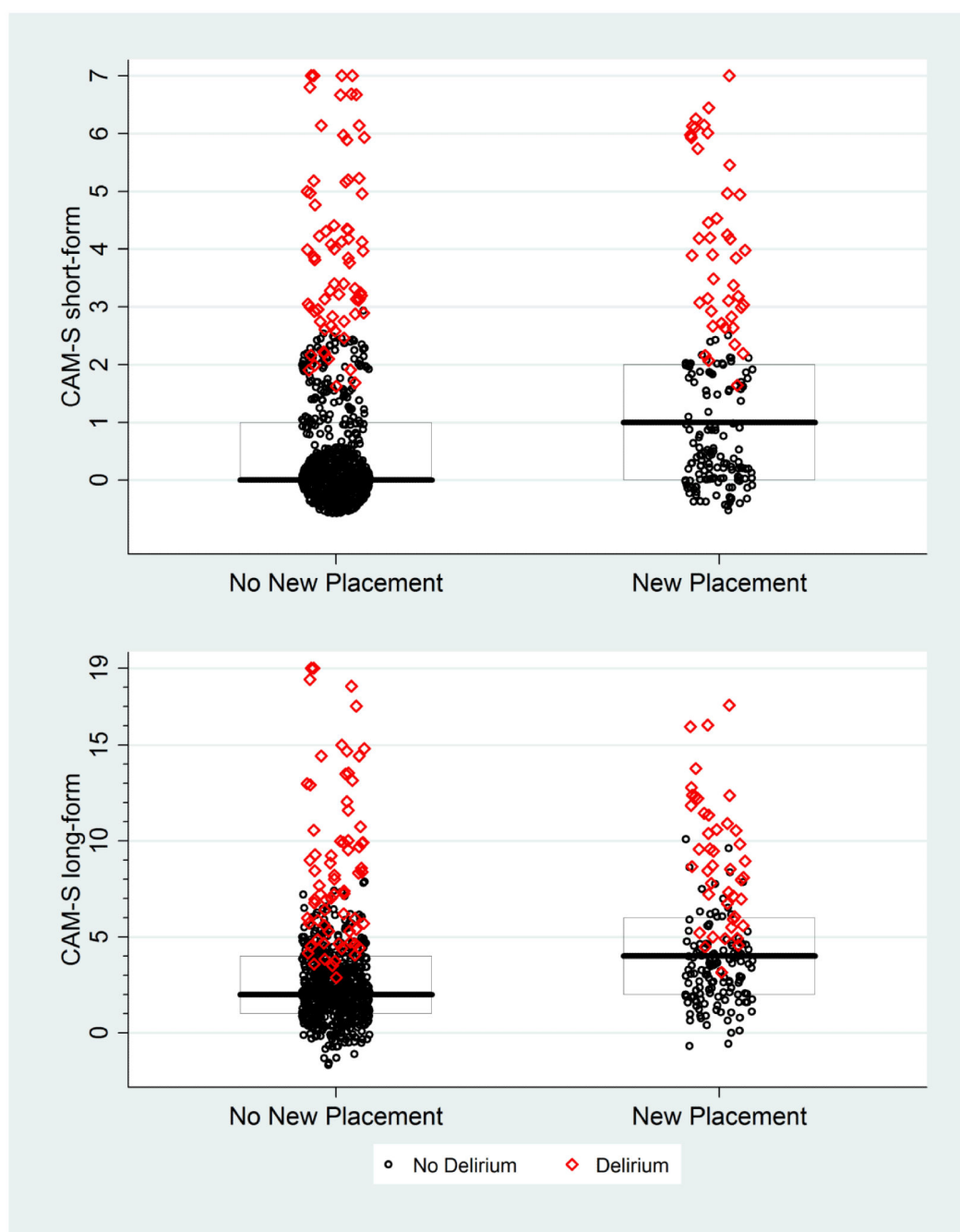
**Appendix Figure 1. CAM-S Scores by Length of Stay\***

\*Plots of maximal CAM-S scores per patient by length of hospital stay. The green line runs through fitted values derived from log-gamma regression.



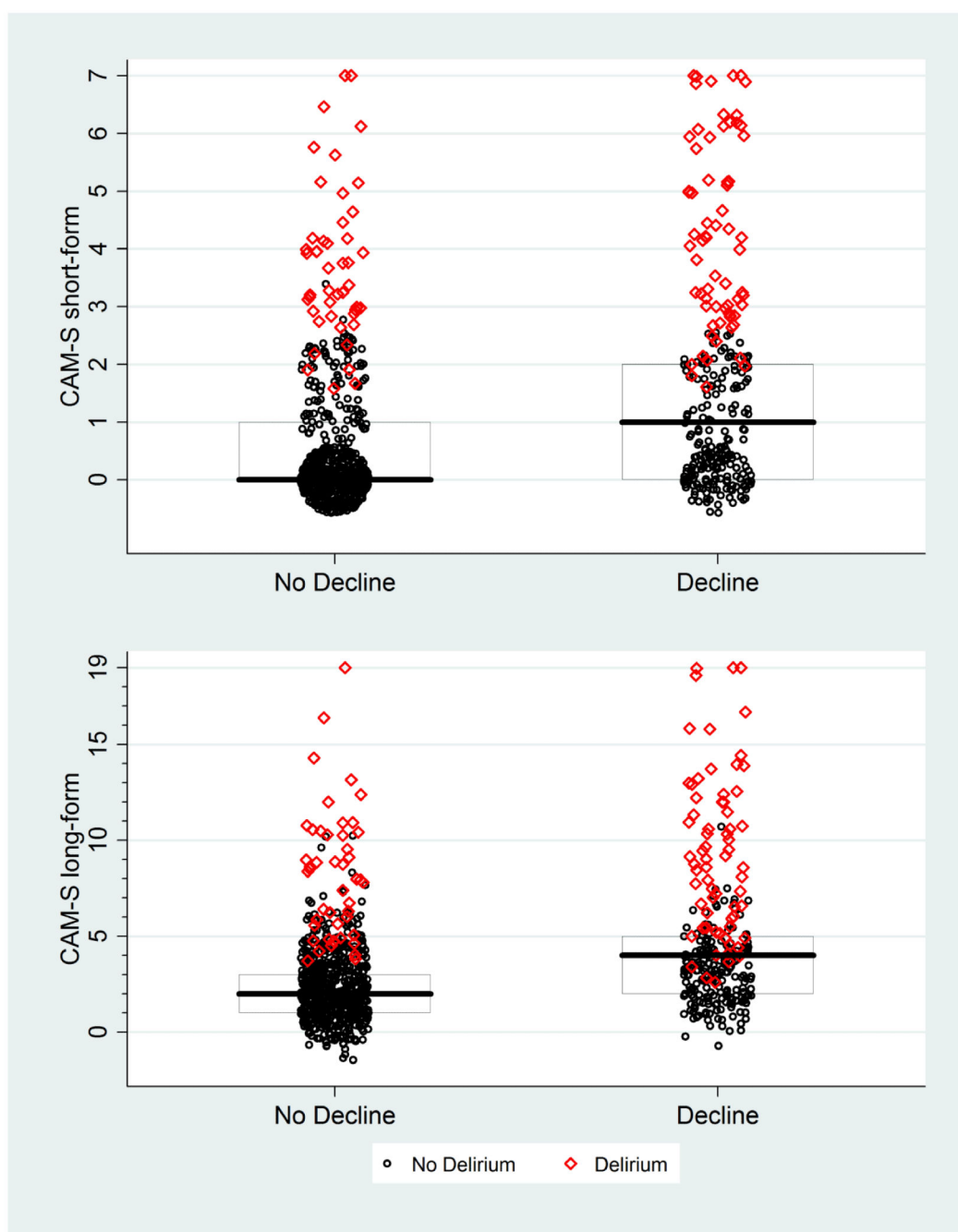
**Appendix Figure 2. CAM-S Scores by Hospital Costs**

\*Plots of maximal CAM-S scores per patient by hospital costs. The green line runs through fitted values derived from log-gamma regression.



**Appendix Figure 3. CAM-S Scores by New Nursing Home Placement\***

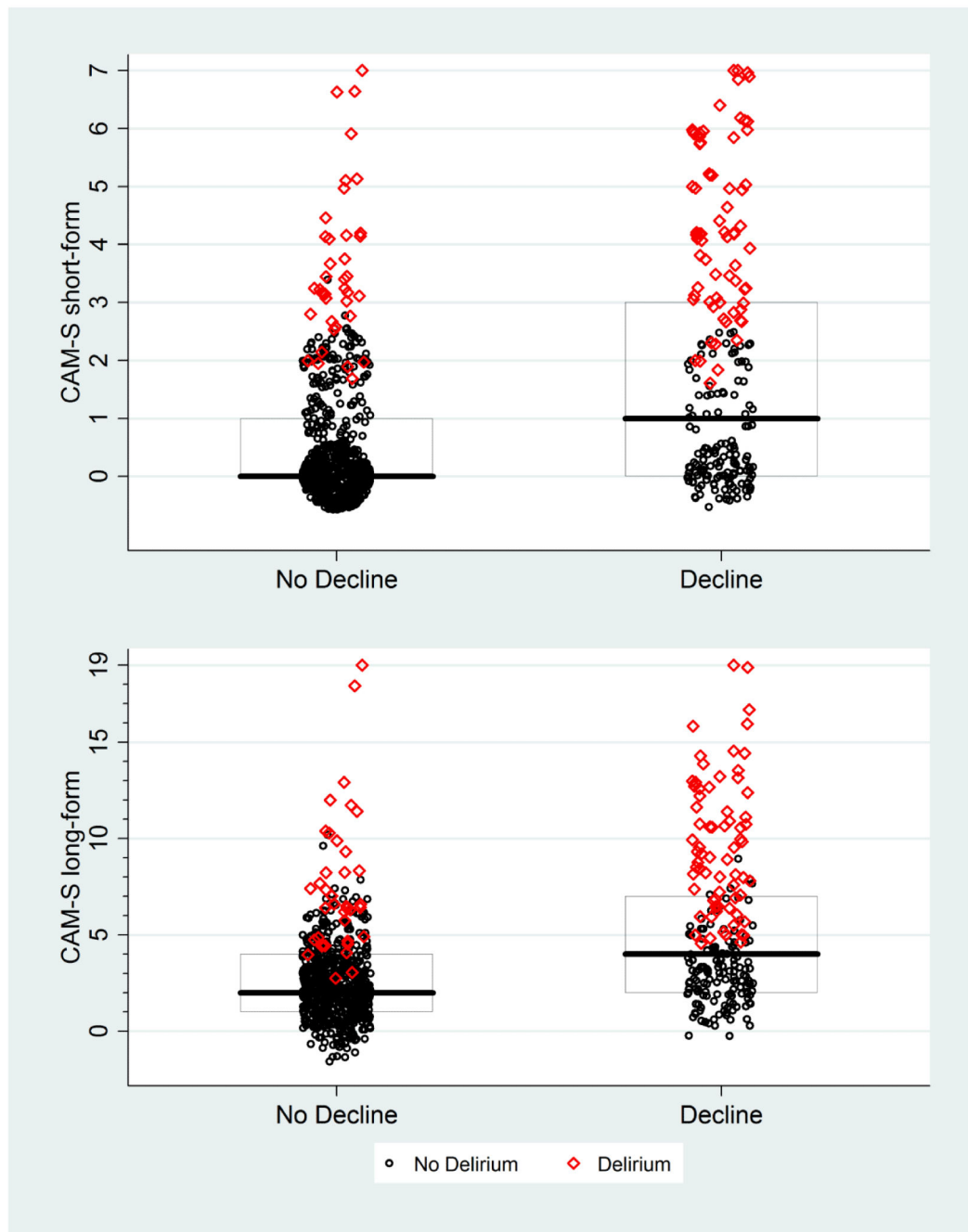
\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.



**Appendix Figure 4. CAM-S Scores by Functional Decline\***

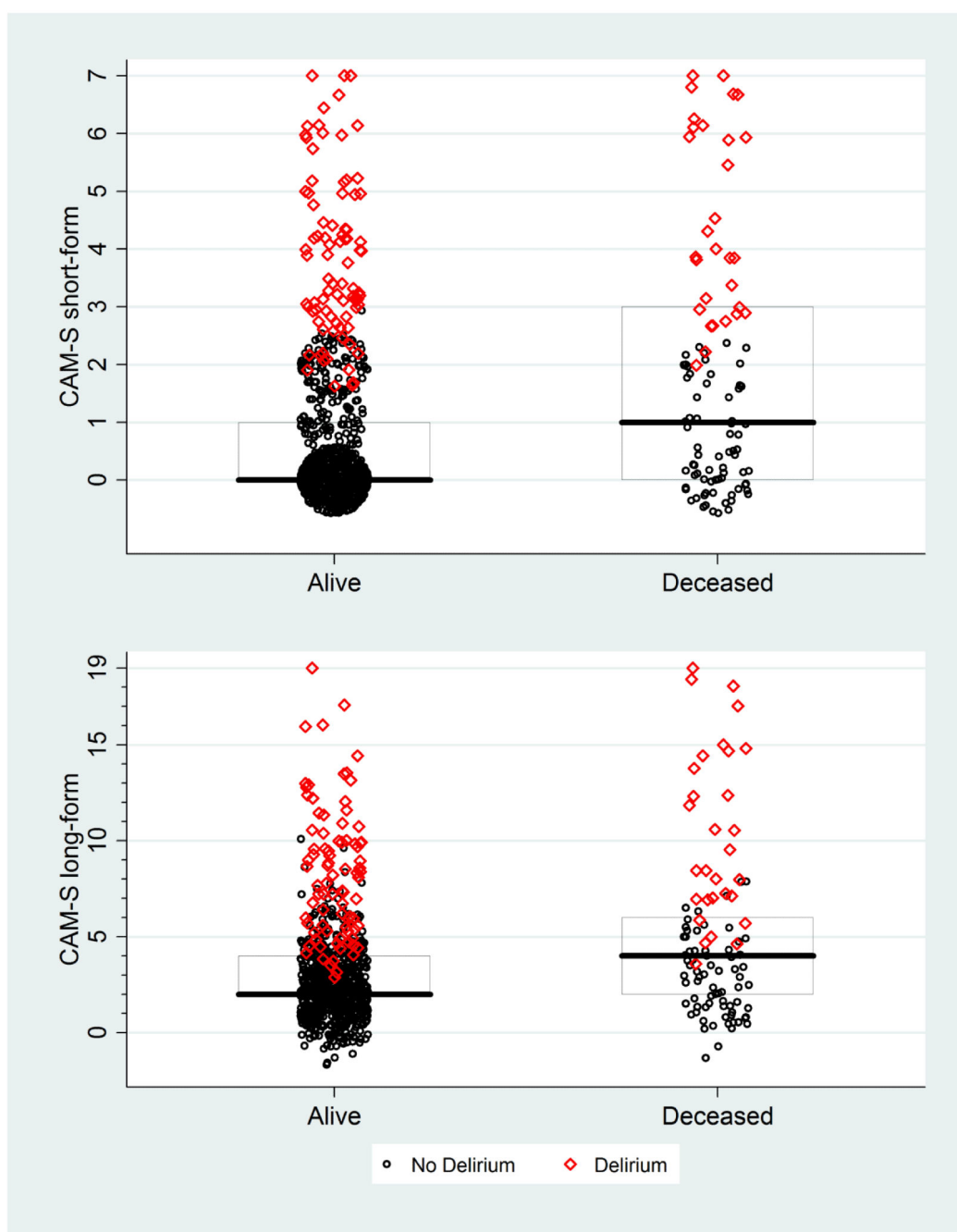
\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Functional decline was defined as a decline of 2 or more points of 14 points in activities of daily living between baseline and discharge scores (see text for details). Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.





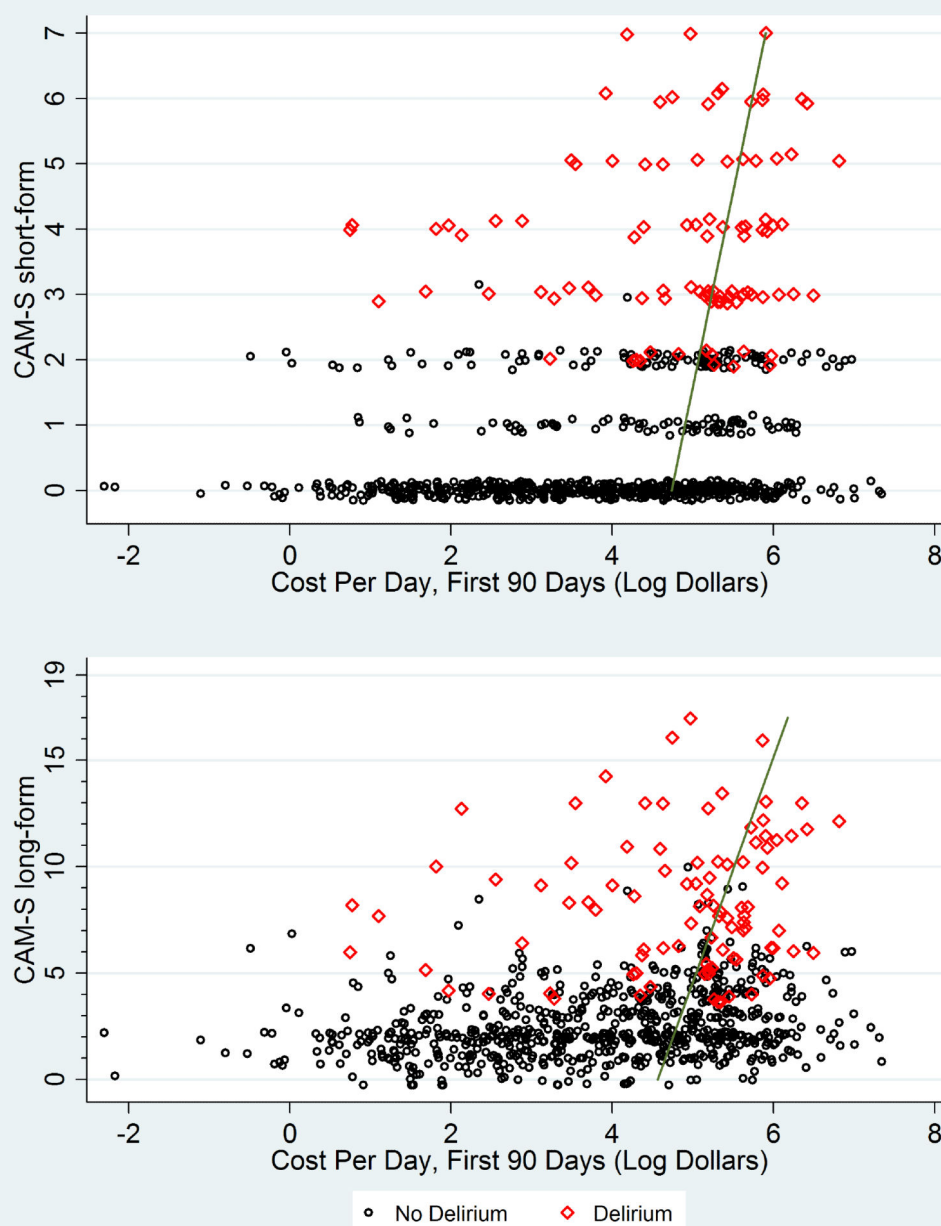
**Appendix Figure 5. CAM-S Scores by Cognitive Decline\***

\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Cognitive decline defined as a decline by 2 or more points in Mini-Mental Status Examination Score between baseline and discharge. Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.



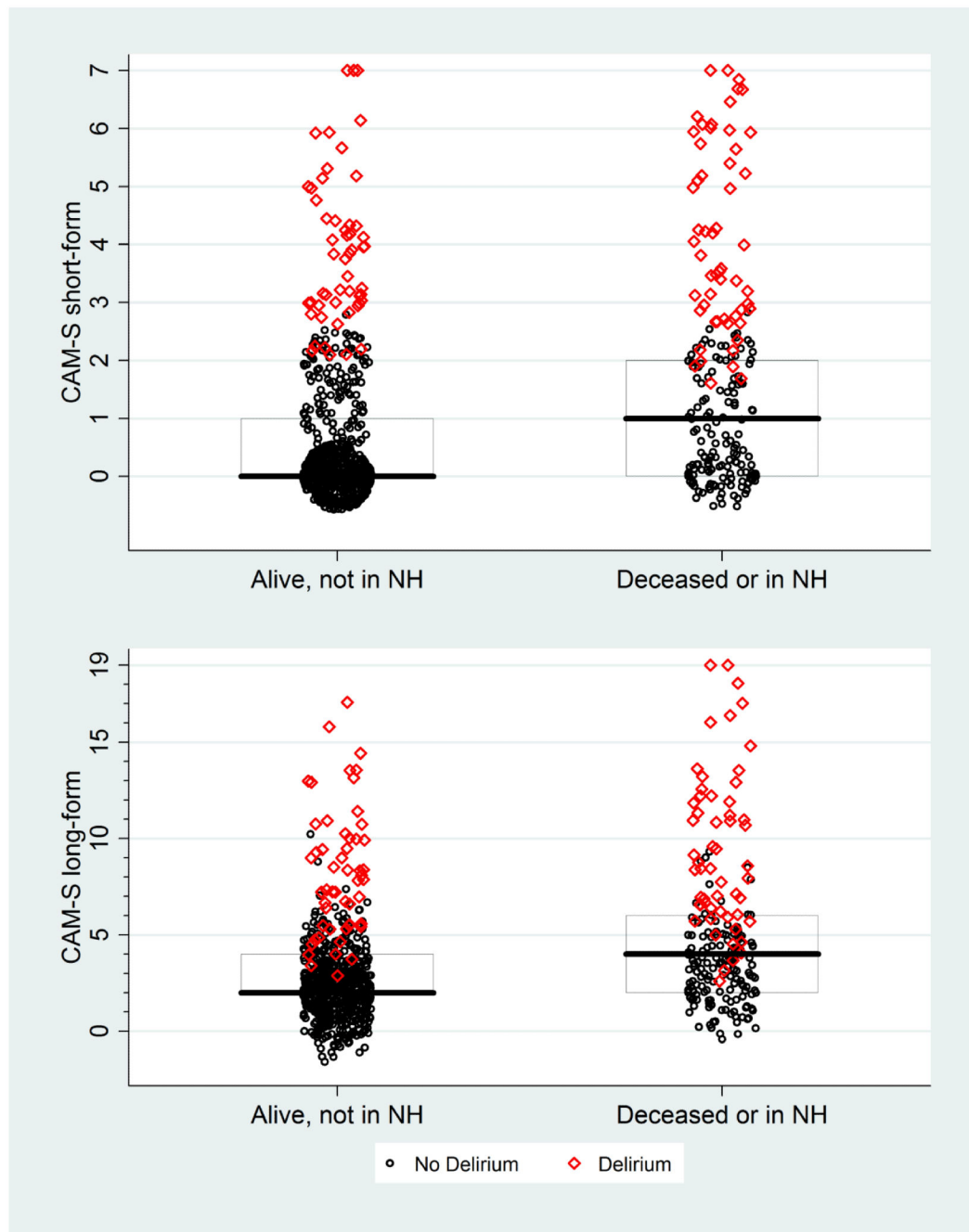
#### Appendix Figure 6. CAM-S Scores by Death within 90 Days

\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.



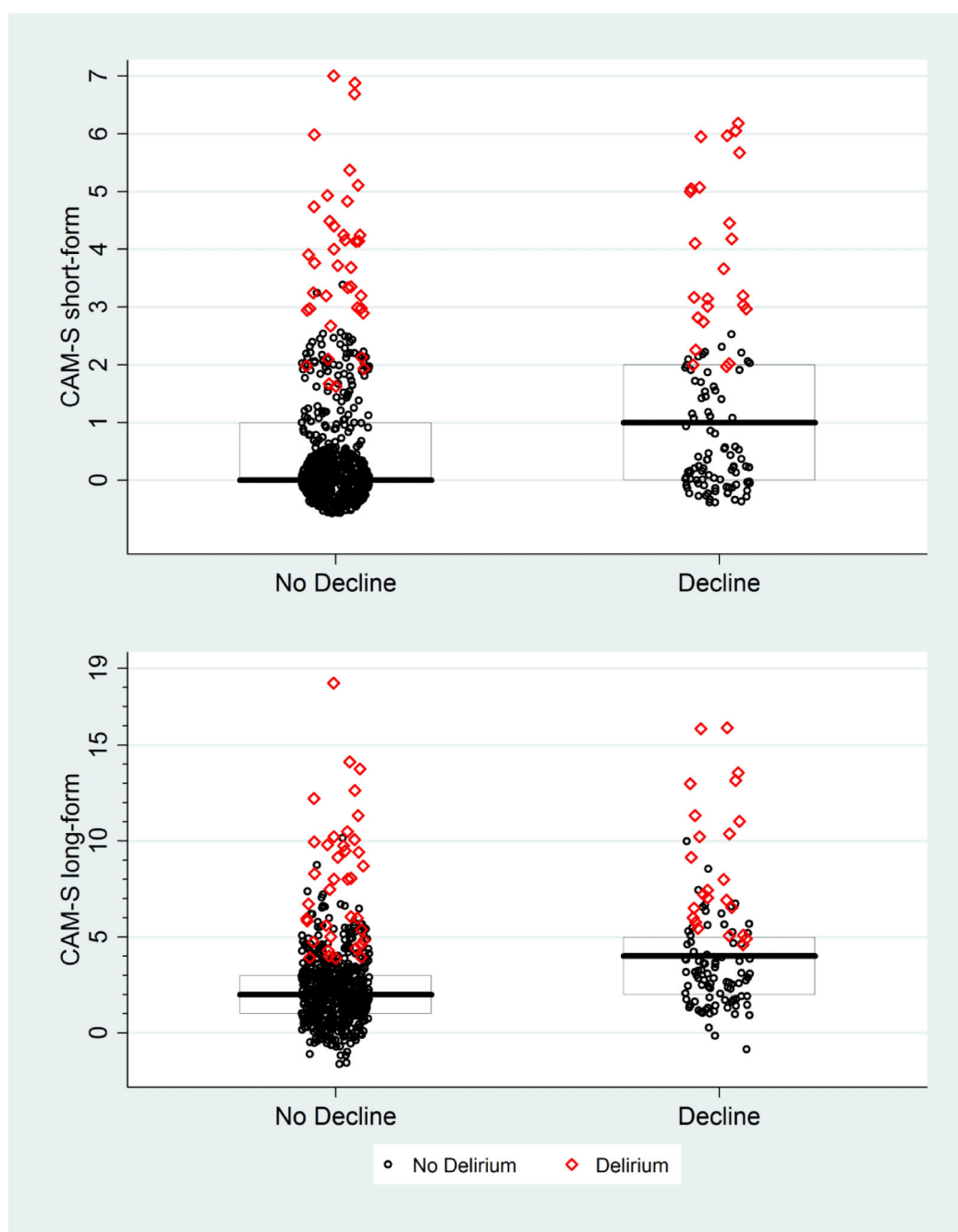
**Appendix Figure 7. CAM-S Scores by Costs Per Day (90 Day Follow-Up)**

\*Plots of maximal CAM-S scores per patient by costs per day. The green line runs through fitted values derived from log-gamma regression.



**Appendix Figure 8. CAM-S Scores by Death or Nursing Home (NH) Residence at 90 Days\***

\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.



**Appendix Figure 9. CAM-S Scores by Functional Decline at 30 Days\***

\* The maximal CAM-S score during each patient's hospitalization was used in all analyses. Functional decline was defined as a decline of 2 or more points of 14 points in activities of daily living between baseline and one-month follow-up scores (see text for details). Boxes around the plots represent the median, 25<sup>th</sup>, and 75<sup>th</sup> percentiles. If a box is not shown, that indicates the median, 25<sup>th</sup> and/or 75<sup>th</sup> percentile all had the same value.

**Table 1**  
**Characteristics of Study Populations**

Characteristic	SAGES Study (N=300)	Project Recovery (N=919)
Age--mean yrs (SD)	76.9 (5.0)	80.0 (6.5)
Male sex--n (%)	134 (45)	365 (40)
Non-white race--n (%)	21 (7)	119 (13)
Married--n (%)	185 (62)	332 (36)
Living alone--n (%)	85 (28)	371 (40)
Residence in nursing home--n (%)	0 (0)	56 (6)
Education--mean yrs (SD)	15 (2.9)	11.1 (3.5)
Charlson comorbidity--mean score (SD)	1.0 (1.3)	2.9 (2.2)
Any impairment in basic activities of daily living--n (%) <sup>*</sup>	21 (7)	320 (35)
Any impairment in instrumental activities of daily living--n (%) <sup>*</sup>	73 (24)	799 (87)
Any cognitive impairment at baseline <sup>†</sup>	24 (8)	406 (44)
Dementia at baseline <sup>‡</sup>	5 (2)	121 (13)
Delirium during hospitalization	68 (23)	115 (13)

SAGES=Successful Aging after Elective Surgery Study; SD=standard deviation.

<sup>\*</sup> Impairment defined as any impairment in one or more basic (or instrumental) activity of daily living

<sup>†</sup> Any cognitive impairment defined by 3MS score < 85 in SAGES and MMSE score <24 in Project Recovery.

<sup>‡</sup> In SAGES, although clinically documented dementia was excluded, early dementia was detected by a Clinical Consensus panel after enrollment in 2% of patients. In Project Recovery, dementia was defined by having either (1) an modified Blessed Dementia Rating Scale (mBDRS) score greater than 4 or (2) an mBDRS score greater than 2, a Mini-Mental State Examination score < 20, and a duration of cognitive symptoms of at least 6 months.

Table 2

The Association of CAM-S with Hospital Outcomes\*

Hospital Outcome	Length of Stay (N=919)	Hospital Costs (N=919)	New Nursing Home Placement (N=851)		Functional Decline (N=908) <sup>†</sup>		Cognitive Decline (N=902) <sup>‡</sup>	
	Adj. Mean, Days (95% CI)	Adj. Mean (95% CI), \$ × 1000	n/N (%)	Adjusted RR (95% CI)	n/N (%)	Adjusted RR (95% CI)	n/N (%)	Adjusted RR (95% CI)
CAM-S Short Form Severity Rating								
None (n=598)	6.5 (6.2, 6.9)	5.1 (4.8, 5.5)	61/571 (11)	Referent	212/589 (36)	Referent	93/589 (16)	Referent
Low (n=91)	8.8 (7.6, 9.9)	7.0 (5.8, 8.2)	14/85 (16)	1.4 (0.8, 2.3)	46/90 (51)	1.4 (1.1, 1.8)	23/89 (26)	1.6 (1.1, 2.4)
Moderate (n=128)	11.1 (9.9, 12.3)	9.7 (8.3, 11.0)	29/112 (26)	2.1 (1.4, 3.1)	80/128 (63)	1.7 (1.4, 2.0)	34/126 (27)	1.6 (1.1, 2.3)
High (n=102)	12.7 (11.2, 14.3)	13.2 (11.1, 15.3)	32/83 (39)	2.5 (1.6, 3.7)	69/101 (68)	1.8 (1.5, 2.2)	64/98 (65)	3.9 (3.0, 5.0)
<i>P</i> <sub>trend</sub>	<.001	<.001	---	<.001	---	<.001	---	<.001
CAM-S Long Form Severity Rating								
None (n=205)	5.6 (5.1, 6.1)	4.2 (3.7, 4.7)	13/198 (7)	Referent	50/202 (25)	Referent	20/202 (10)	Referent
Low (n=288)	6.8 (6.4, 7.3)	5.4 (4.9, 5.9)	25/278 (9)	1.4 (0.7, 2.5)	115/284 (40)	1.6 (1.2, 2.1)	53/282 (19)	1.9 (1.1, 3.0)
Moderate (n=234)	8.8 (8.1, 9.5)	7.3 (6.5, 8.1)	40/215 (19)	2.3 (1.3, 4.2)	126/231 (55)	2.1 (1.6, 2.8)	47/230 (20)	2.0 (1.2, 3.3)
High (n=192)	11.9 (10.8, 12.9)	11.4 (10.0, 12.8)	58/160 (36)	3.9 (2.1, 7.0)	116/191 (61)	2.3 (1.8, 3.0)	94/188 (50)	4.8 (3.0, 7.7)
<i>P</i> <sub>trend</sub>	<.001	<.001	---	<.001	---	<.001	---	<.001

Adj=adjusted; RR=relative risk; CI=confidence interval. Short Form, points: None=0, Mild=1, Moderate=2, Severe=3-4, Severe=5-19. All models are adjusted for age, sex, nonwhite race, Apache II score, Charlson comorbidity index, and baseline dementia. All models except the model for functional decline were also adjusted for baseline impairment in activities of daily living. Nursing home analyses excluded 54 patients residing in nursing home at baseline and 14 who died during hospitalization. The functional decline analyses excluded 11 patients with missing ADL information at discharge; the cognitive decline analyses excluded 17 patients with missing MMSE values at discharge.

\* Analyses conducted in Project Recovery. The maximal CAM-S score during each patient's hospitalization was used in all analyses.

<sup>†</sup> Functional decline defined as decline by 2 or more points on the 14-point Activities of Daily Living (ADL) score (equivalent to 1 ADL on the standard 7-point scale) between baseline and discharge.

<sup>‡</sup> Cognitive decline defined as decline by 2 or more points on the 30-point Mini-Mental State Examination (MMSE) score between baseline and discharge.



Table 3

The Association of CAM-S with Post-Hospital Outcomes\*

Post-Hospital Outcome	Death within 90 Days (N=919)		Cost Per Day, First 90 Days (N=831)		Death or Nursing Home Residence at 90 Days (N=844)		Functional Decline at 30 Days (N=712) <sup>†</sup>	
	n/N (%)	Adjusted RR (95% CI)	Adj. Mean (95% CI), \$	n/N (%)	Adjusted RR (95% CI)	n/N (%)	Adjusted RR (95% CI)	
CAM-S Short Form Severity Rating								
None (n=598)	39/598 (7)	Referent	115.8 (100.2, 131.3)	81/544 (15)	Referent	143/497 (29)	Referent	
Low (n=91)	14/91 (15)	2.0 (1.1, 3.5)	158.4 (108.0, 208.7)	27/82 (33)	1.9 (1.3, 2.7)	29/70 (41)	1.4 (1.0, 1.9)	
Moderate (n=128)	20/128 (16)	1.8 (1.1, 3.2)	175.3 (129.8, 220.8)	48/121 (40)	2.1 (1.5, 2.9)	41/91 (45)	1.5 (1.2, 2.0)	
High (n=102)	28/102 (27)	3.3 (2.1, 5.1)	168.2 (116.4, 220.0)	49/97 (51)	2.5 (1.9, 3.3)	28/54 (52)	1.6 (1.2, 2.2)	
<i>P</i> <sub>trend</sub>	---	<.001	<.001	---	<.001	---	<.001	
CAM-S Long Form Severity Rating								
None (n=205)	14/205 (7)	Referent	97.6 (76.0, 119.2)	24/181 (13)	Referent	35/175 (20)	Referent	
Low (n=288)	20/288 (7)	0.8 (0.4, 1.5)	126.7 (103.8, 149.7)	37/263 (14)	1.0 (0.6, 1.6)	67/238 (28)	1.4 (1.0, 2.0)	
Moderate (n=234)	25/234 (11)	1.3 (0.7, 2.4)	138.0 (110.2, 165.7)	57/217 (26)	1.6 (1.0, 2.5)	77/180 (43)	2.0 (1.4, 2.8)	
High (n=192)	42/192 (22)	2.3 (1.3, 4.1)	177.2 (138.2, 216.3)	87/183 (48)	2.5 (1.6, 3.7)	62/119 (52)	2.3 (1.6, 3.3)	
<i>P</i> <sub>trend</sub>	---	<.001	<.001	---	<.001	---	<.001	

RR=relative risk; CI=confidence interval. Short Form, points: None=0, Mild=1, Moderate=2, Severe=3-7; Long Form, points: None = 0-1, Mild=2, Moderate=3-4, Severe=5-19. Deaths within 90 days include all in-hospital deaths. All models are adjusted for age, sex, nonwhite race, Apache II score, Charlson comorbidity index, and baseline dementia. All models except the model for functional decline were also adjusted for baseline impairment in activities of daily living, Medicare data (nursing home and costs) were missing for 75 patients (including those receiving care in health maintenance organizations). The cost per day analyses additionally excluded 13 patients who died during hospitalization. See text for details.

\* Analyses conducted in Project Recovery. The maximal CAM-S score during each patient's hospitalization was used in all analyses.

<sup>†</sup> Functional decline defined as decline by 2 or more points on the 14-point Activities of Daily Living (ADL) score (equivalent to 1 ADL on the standard 7-point scale) between baseline and 30 days. These analyses included all 728 patients who were available for telephone follow up interviews at one-month, but excluded 16 with missing ADL data.